and forty-three respectively. In the fourth grade group, there was no observer who found the merely incongruous picture the funniest completion in more than four cases out of the possible ten; there were two who found it so in four cases, five who found it so in three cases, eight who found it so in two cases, three who found it so in one case. In the seventh grade, there were three observers who found the merely incongruous picture funniest in six cases out of the ten; two who found it so in five cases, four in four cases, four in three cases, three in two cases, no one who found the incongruous picture funniest in only one or in no case. In the first group of college girls, two found the incongruous picture funniest in six cases, no one in five, three in four cases, two in three cases, four in one case, two never. In Group II, college girls, no one found the incongruous picture funniest in six cases, two so found it in five cases, three in four cases, two in three cases, four in two cases, four in one case, two never. In Group III, college girls, one observer found the merely incongruous picture funniest in seven cases, no one in six cases, one in five cases, two in four cases, three in three cases, six in two cases, two in one case, two never. In Group IV, college girls, one observer found the incongruous picture funniest in seven cases, no one in six cases, two in five cases, one in four cases, four in three cases, four in two cases, two in one case, three never.

It will be seen that there is much more individual variation in the taste for the 'purely incongruous' style of humor among the adults than among the children. It is obviously desirable to extend the experiment to much larger numbers of fourth and seventh grade children, and this we plan to do.

XL. The Results of Certain Standard Mental Tests as Related to the Academic Records of College Seniors

By Hermine Baum, Miriam Litchfield, and M. F. Washburn

The tests used were the following:

(1) The Hard Opposites Test of Woodworth and Wells. The results were stated in terms of the average time for a single correct response.

(2) The Analogies Test, lists A and B, of the Woodworth and Wells set: the results were stated in terms of the mean between the average time for a single correct response in A and that for a single correct response in B.

(3) The Substitution Test, the Woodworth and Wells form, with the star, circle, square, cross, and triangle. The results were stated in terms of an index of efficiency, accuracy being reckoned by subtracting for each error two from a total of 100, and the time required for the whole sheet being divided by the index of accuracy thus obtained.

(4) The Cancellation Test, the form beginning hplg or that beginning zcyu , a's being cancelled. The results were stated in the form of the index of efficiency, for the work-limit method: 100 times the index of accuracy divided by the time; the index of accuracy being found by dividing the number of a's crossed by the number on the page.

(5) The Information Test. This was the familiar Whipple list of words, beginning 'ageratum.' Since an observer may often think he can define a word when he cannot, and his opinion can be refuted only by a laborious criticism of definitions, we took as the measure of an individual's information the number of words on this list that were marked N, as being wholly new and unfamiliar to her. One is not likely to be mistaken on such a point as this.

No preliminary practice was given on any of the tests except the Opposites and the Analogies Tests, where examples were given.

The observers were chosen on the following principle. They were all Vassar students in the second semester of their senior year. At this period the Dean's Office has on record for each girl a numerical value which represents both the quantity and the quality of her work for the seven semesters she has completed. For each semester hour of work done of grade A, she receives 5 points of credit; each hour of B grade work gives her three points, each hour of C work 2 points; each hour of D work (just passing) one point. The object of our study was to see how well the work of a student, thus measured, correlated with her performance in the above tests.

The study may be divided into two parts.

I. Thirty-eight seniors were chosen whose records formed a fairly continuous scale from the highest to the lowest in the class. The plan was to have their numerical standings differ by four points. For the lowest twenty-three on the scale, this was actually the case; in the upper half of the scale there was occasionally a greater distance between the ranking of one individual and the next above or below her. Beside this ranking as regards academic performance we set the rankings of the students in the various tests, and calculated rank difference correlations. The results were as follows.

Index of	Correlation	between	Cancellation	and	Academic	Rank,	+ .06,	P.E.	.11.
"	"	"	Substitution	"	"	"	+ .03,	P.E.	.11.
"	"	"	Opposites	"	"	"	+ .30,	P.E.	.09.
"	"	"	Analogies	"	"	"	+ .39,	P.E.	.09.
"	"	"	Analogies and Opposites combined	"	"	"	+ .40,	P.E.	.09.
"	"	"	Information	"	"	"	+ .33,	P.E.	.09.

Considering the very slight differences in academic rank on which the scale of academic ranks was based, the Analogies, Opposites, and Information Tests make a good showing.

II. Two groups of twenty-five students each were chosen. The academic ranks of all in group I were represented by numerical values over 350; the ranks of those in Group II were represented by numerical values under 210.

The results of the tests for these two groups, I containing the seniors with best records, II those with poorest records, were as follows:

Cancellation: average index of efficiency, Group 1, .69, m.v. 11.

Cancellation: average index of efficiency, Group 11, .74, m.v. 15.

The group of lower academic standing did somewhat better than that of higher academic standing. The distribution of individual scores confirms this: the numbers were as follows. Above 1, in Group I, 1; in Group II, 1. Between .9 and 1, Group I, 5; Group II, 3. Between .8 and .9, Group I, 3; Group II, 6. Between .7 and .8, Group I, 4; Group II, 7. Between .6 and .7, Group I, 10; Group II, 4. Between .5 and .6, Group I, 2; Group II, 3. Below .5, Group I, none; Group II, 1.

Substitution: average index of efficiency, Group I, 1.31, m.v. .13.
Substitution: average index of efficiency, Group II, 1.34, m.v. .17.

The smaller this index, the better the performance. Here the group of higher academic standing did a very little better. The distribution of individual scores is in accordance: The number of scores above 1.7 was, for Group I, none; for Group II, 1, Scores between 1.6 and 1.7 were made by 2 observers in Group I; by 3 in Group II. Scores between 1.5 and 1.6 were made by 4 in Group I and by 3 in Group II. Scores between 1.4 and 1.5 were made by 1 in Group I and by 5 in Group II. Scores between 1.3 and 1.4 were made by 6 in Group I and by 2 in Group II. Scores between 1.2 and 1.3 were made by 6 in Group I and by 3 in Group II. Scores between 1.1 and 1.2 were made by 4 in Group I and by 6 in Group II. Scores between 1 and 1.1 were made by 2 in Group I and by none in Group II. Scores between .9 and 1 were made by none in Group I and by 2 in Group II.

Opposites. In this test the smaller the score the better the performance.

Average time of a correct response, Group I, 2.87, m.v. .73.
Average time of a correct response, Group II, 3.58, m.v. 1.37.

Here the group of superior academic standing does distinctly better than that of poorer academic standing, and the distribution of cases bears this out. Scores of between 1 and 2 seconds, that is, the best scores, were made by four from Group I and by 2 from Group II. Scores between 2 and 3 seconds were made by 16 from Group I and by 7 from Group II. Scores between 3 and 4 seconds were made by 4 from Group I and by 7 from Group II. Scores between 4 and 5 seconds were made by 1 from Group II and by 2 from Group 2. Scores above five seconds were made by no one in Group I, and by 7 persons from Group II.

Analogies. Here also the smaller the score the better the performance.

Average time of a correct response, Group I, 2.46, m.v. .26.
Average time of a correct response, Group II, 2.62, m.v. .30.

The better group, academically, is better in the test than the poorer group academically, though the difference is not so marked as in the case of the Opposites Test. The distribution of scores is very scattering, and may be summarized by saying that scores better than 2.5 seconds were made by 56% of the observers in Group I and by 36% of observers in Group II; while scores poorer than 3 seconds were made by 12% of Group I and by 20% of Group II.

Information. Since this test was scored by the number of words new to the observer, the smaller numbers indicate the better performances.

Average number of new words, Group I, 31, m.v. .8.
Average number of new words, Group II, 37, m.v. .78.

An advantage on the part of the superior group in academic standing appears. Fifty-two per cent of the observers in Group I found less than 30 new words; 28% of Group II found less than 30 new words. On the other hand 36% of Group I and 40% of Group II found more than 40 new words, so the difference between the groups seemed to be that there were fewer noticeably good vocabularies in the poorer academic group, but about the same number of noticeably poor vocabularies in the two groups.

The net result of the study is that the Cancellation Test and the Substitution Test do not serve at all to differentiate the best from the poorest students in a group all of whom are up to the standard re-

quired for graduation from college; and that while neither the Opposites Test nor the Analogies Test serves entirely to separate such groups, there is a distinct correlation in the case of both between test performance and academic record. The number of words in the Whipple Information Test which are new to a student is correlated with her academic record: the best records in this test are largely made by students of the best academic standing, but some noticeably poor records are made by such students.

A Note on the Terman Superior Adult Tests, as Applied to Vassar Freshmen

By M. F. Washburn

In the fall of 1917, when the members of the incoming Freshman class were tested, they were put through the Terman Superior Adult Tests. Out of 317 students tested, 46, or 14.5 %, passed all these tests. On 35 of these students there were obtained from their instructors during the year 69 judgments of general ability, made on the Miner Scale, 5 being used to designate the highest and 1 the lowest grade, while 3 indicated average ability. The average of these 69 judgments was 3.36. Eleven of the forty-six had no judgments made upon them. It would appear that our Freshmen 'superior adults' failed to impress their instructors with their superiority, since the average judgment assigns to them about normal ability.